DCPAPERS

Creativity and Authorship in the Age of Artificial Intelligence: A Metadata Perspective

Lala Hajibayova[1,*]

[1]Kent State University, Kent, USA

Abstract

DCMI conceptualization of creator in the context of AI-generated content introduces new complexities, as AI systems challenge traditional notions of authorship, responsibility, and intellectual agency. Large language models, in particular, can generate content autonomously, yet they lack the agency, intent, and legal or moral responsibility that are traditionally associated with a creator or author. Consequently, assigning a 'creator' under this definition complicates questions of accountability, ownership, and provenance within metadata standards. This paper calls for a critical reexamination of how the concept of 'creator' is defined and applied in metadata schemas in the age of generative AI.

Keywords

Artificial Intelligence, Authorship, DCMI

1. Introduction

The rapid advancement of Artificial Intelligence (AI) and its ability to replicate human behavior and interactions has led to a surge of AI-generated content, making it increasingly difficult to distinguish between content created by humans and that produced by AI.

Attribution of authorship for electronic resources has long been considered challenging due to their often fluid and ephemeral nature [1]. The Dublin Core Metadata Initiative (DCMI) defines a creator as "an entity primarily responsible for making the resource" [2]. The DCMI conceptualization of creator in the context of AI-generated content introduces additional complexities, as AI systems challenge traditional notions of authorship, responsibility, and intellectual agency. AI systems, particularly large language models, can generate content autonomously, yet they lack agency, intent, and legal or moral responsibility—key attributes traditionally associated with creator or authorship. As a result, attributing creator under this definition complicates issues of accountability, ownership, and provenance in metadata standards, calling for a reexamination of how concept of creator is conceptualized in metadata schemas in the age of generative AI.

In light of generative AI, this paper advocates for a critical reexamination of the conceptualization and implementation of 'creator' in metadata standards.

DCMI-2025 International Conference on Dublin Core and Metadata Applications
*Corresponding author.
lhajibay@kent.edu (L. Hajibayova)
0000-0001-8358-9858 (L. Hajibayova)

https://doi.org/10.23106/dcmi.952528171

1.1. Attribution of Authorship

Historically, the concept of authorship has been foundational to bibliographic description and regarded as a key part of the main entry, which includes author's name, the title, and the imprint. The main entry was defined in ALA Catalog Rules (1941) as: A full catalog entry, usually the author entry, giving all the information necessary to complete identification of a work. This entry bears also the tracing of all the other headings under which the work in question is entered in the catalog. Charles Cutter defined "main entry" as: the full or principal entry; usually the author entry (Cwiok, 2005). The official definition of main entry has changed surprisingly little over the years.

The crediting of authorship has long followed the so-called 'rule of three,' which involved the listing of only the first three authors and omitting the rest, which was indicated by an ellipsis and, in brackets, a phrase such as [and others], [et d'autres], etc., according to the language of the title page [3].

The DC schema was originally created to provide a framework for describing electronic resources, recognizing their unpredictable and volatile nature. Its elements, including Creator, are optional, repeatable, and extensible, thereby enhancing the flexibility of the core set of descriptors. However, the broad and loosely defined nature of the Creator element presents challenges in its consistent application. There is also significant overlap in meaning between the Creator, Contributor, and sometimes Publisher elements in Dublin Core [2], further challenging differentiation of the responsibilities:

Creator: An entity primarily responsible for making the content of the resource.

Publisher: An entity responsible for making the resource available.

Contributor: An entity responsible for making contributions to the resource.

The schemas that developed based on DCMI share many of the same issues surrounding Creator element attribution. For example, Visual Resources Association Core (VRA) and Gateway to Educational Materials (GEM) [1].

VRA Core: Modeled on Dublin Core to describe visual works and their surrogates. It supports the one-to-one principle, designating separate records for works and images. This means an original object and its digital surrogate could have different creators. VRA provides a more specific framework for responsibility with qualifiers for the Creator element: Creator.Role, Creator.Attribution, Creator.Personal name, and Creator.Corporate name. This approach makes responsibility attribution less ambiguous than in Dublin Core [1, 4].

GEM: Modeled after Dublin Core, adding GEM-specific elements and qualifiers. GEM extends DC by adding a Role qualifier to Creator, Contributor, and Publisher elements to help implementers specify roles. However, this still raises questions about intellectual responsibility, authorship, and accurate resource description, as it's unclear who is the "creator" when considering original authors, webmasters, or compilers [1].

Despite modifications like qualifiers, overlap and ambiguity in defining the Creator element persist in these schemas. A fundamental difficulty lies in determining what the metadata schema should describe – the digital object "in hand" or the original work it represents – and what purpose the Creator element should serve. While users might be interested in the original creator (e.g., Picasso for an image), the information about the entity that created the digital surrogate is also important [1]. Overall, the schemas that do not include the term "Creator"

https://doi.org/10.23106/dcmi.952528171

tend to be more complex and specific in their element sets, which helps reduce ambiguity in attributing responsibility [1].

However, this approach to the authorship principle, i.e., idea that works are primarily identified through their creators, is challenged in collaborative, fluid, and often ephemeral nature of electronic resources, which disrupts these traditional notions of authorship and bibliographic control. The rise of digital and AI-generated content further complicates authorship attribution, prompting a reevaluation of established frameworks [5]. The ambiguity surrounding the Creator element in metadata standards such as Dublin Core reflects this ongoing shift away from conventional authorship models in the digital age, highlighting the need for updated conceptual and descriptive approaches. The difficulty in applying these schemas to representation of resources implies a need to reassess how "authorship" is attributed- shifting focus from "who is author/creator of the content?" to "who is responsible for this content" and "what is the origin of the content" [6]. The problem with the Creator element might suggest a move from focusing on "authorship" and focusing on "intellectual responsibility." The critical assessment of the resources requires a shift in focus from simple attribution of the creator to the origination of content and responsibility for the content. A potential model for refining the Creator element can be found in the Text Encoding Initiative (TEI) guidelines, which offer detailed mechanisms for representing roles and responsibilities in textual production. TEI distinguishes clearly between entities responsible for authorship, editing, annotation, and other forms of contribution. Adopting a similar level of specificity in metadata schemas would allow cultural heritage institutions and knowledge organization systems to more accurately delineate responsibilities, thereby supporting better provenance tracking and ethical attribution practices in contexts involving both human and machine contributions.

To address the growing complexities introduced by AI-generated content and collaborative forms of authorship, it is essential to refine the definition of the Creator element in metadata schemas. Traditional definitions often presume a human-originated source of authorship, but this assumption no longer holds in environments where generative AI systems contribute to or fully produce content. A more nuanced approach is needed, one that attributes responsibility more accurately and reflects the multiplicity of actors involved in content creation. This approach is essential to facilitate the accurate and authoritative representation of resources within knowledge organization systems. As digital content increasingly emerges from diverse sources, including collaborative human efforts, automated AI systems, and hybrid interactions, traditional attribution mechanisms struggle to convey the full context of authorship and responsibility. Enhancing metadata to reflect these distinctions ensures that resources are described with greater accuracy, which is critical for maintaining the trustworthiness, provenance, and scholarly value of the knowledge. Accurate representation not only upholds the integrity of collections but also supports informed access, ethical reuse, and responsible knowledge dissemination in an era of rapidly evolving content creation.

One strategy involves introducing a sub-element, such as `<creator.origin>`, which would specifically capture the original source or intellectual provenance of the content. This sub-element could be used to denote whether the content originated from a human author, an AI system, or a collaborative effort. It would allow metadata records to differentiate between human-created works, AI-assisted outputs, and content wholly generated by machine learning models. This level of granularity would provide greater clarity about the nature of the resource

https://doi.org/10.23106/dcmi.952528171

and facilitate more accurate attribution, copyrights, and ethical accountability.

Moreover, implementing such distinctions would align metadata practices with principles of transparency, traceability, and responsible information stewardship. It would also support downstream uses of metadata, such as in digital preservation, citation practices, and provenance research, by allowing users and systems to make informed decisions about the credibility, originality, and ethical considerations of the resource. Ultimately, refining the Creator element and incorporating sub-elements like can serve as a critical step in adapting knowledge organization systems to the realities of AI-mediated content creation, while preserving the integrity and trustworthiness of cultural and scholarly records.

2. Conclusions

Cultural heritage institutions, such as museums, libraries, and archives, have traditionally functioned as stewards of society's collective memory, preserving works regarded as accurate and authoritative accounts of human history, culture, and knowledge. Upholding high standards of authenticity and accuracy is therefore essential to maintaining the integrity of this preserved heritage. DCMI schema should facilitate not only differentiate AI-generated content from human-authored work but also adopt ethical, human-centered principles to ensure the quality, accuracy, and responsible representation of knowledge from all sources. This calls for a reimagined approach to metadata design, including the development of more nuanced and clearly defined elements that distinguish between creators and contributors. Such refinement is necessary to account for the diverse roles involved in content production, especially in contexts where human and AI contributions intersect and to support more transparent attribution, accountability, and provenance within knowledge organization systems.

References

[1] Cwiok, J. (2005). The Defining Element – A Discussion of the Creator Element within Metadata Schemas. In R. Smiraglia (Ed.), *Metadata: A Cataloger's Primer*. Routledge. doi: https://doi.org/10.4324/9780203051238

[2] DCMI (n.d.). https://www.dublincore.org/specifications/dublin-core/dcmi-terms/

[3] Howarth, L. C., & Weihs, J. (2008). Enigma Variations: Parsing the Riddle of Main Entry and the "Rule of Three" from AACR2 to RDA. *Cataloging & Classification Quarterly*, 46(2), 201–220. https://doi.org/10.1080/01639370802177620

[4] Martynovich, S. (2024). Toward an extended metadata standard for digital art. *Journal of Documentation*, 80(2), 469–486. https://doi.org/10.1108/JD-07-2023-0126

[5] Huang, B., Chen, C., & Shu, K. (2025). Authorship attribution in the era of LLMs: Problems, methodologies, and challenges. *ACM SIGKDD Explorations Newsletter*, 26(2), 21-43.

[6] Hajibayova, L., & Lee, S. (2025). Authorship in the Age of Artificial Intelligence. *Journal of Documentation*.

https://doi.org/10.23106/dcmi.952528171